Appl Math Optim 29:161–186 (1994)

Applied Mathematics & Optimization

Local Analysis of Newton-Type Methods for Variational Inequalities and Nonlinear Programming

J. Frédéric Bonnans

INRIA, B.P. 105, 78153 Rocquencourt, France

Communicated by J. Stoer

Abstract. This paper presents some new results in the theory of Newton-type methods for variational inequalities, and their application to nonlinear programming. A condition of semistability is shown to ensure the quadratic convergence of Newton's method and the superlinear convergence of some quasi-Newton algorithms, provided the sequence defined by the algorithm exists and converges. A partial extension of these results to nonsmooth functions is given. The second part of the paper considers some particular variational inequalities with unknowns (x, λ), generalizing optimality systems. Here only the question of superlinear convergence of $\{x^k\}$ is considered. Some necessary or sufficient conditions are given. Applied to some quasi-Newton algorithms they allow us to obtain the superlinear convergence of $\{x^k\}$. Application of the previous results to nonlinear programming allows us to strengthen the known results, the main point being a characterization of the superlinear convergence of $\{x^k\}$ assuming a weak second-order condition without strict complementarity.

Key Words. Variational inequalities, Nonlinear programming, Successive quadratic programming, Superlinear convergence.

AMS Classification. Primary 90C30, Secondary 65K05, 90C26.

1. Introduction

This paper is devoted to the local study of Newton-type algorithms for variational inequalities. Variational inequalities have been studied for a long time (see [16]) mainly because of their applications to mechanical systems. The operators in that field are often monotone, and a large theory of monotone operators has been developed (see [6]); several algorithms for convex programming, including duality methods, have been extended to this framework (see [11]). Some problems in economy as well as optimality systems of nonlinear programming problems can also be represented by variational inequalities (see [21] and [13]). Consequently, the strength and large use of Newton-type algorithms for nonlinear programming, the so-called successive quadratic programming (see [2] and [10]), suggests developing a theory of Newton-type methods for variational inequalities (we do not speak here of some different approaches of Newton-type algorithms for variational inequalities—reviewed in the survey by Harker and Pang [13]). Some early (but unpublished) works in this direction due to Josephy [14], [15] give a local analysis using the concept of strong regularity [19]. Josephy obtains a quadratic rate of convergence for Newton's method and superlinear convergence for some quasi-Newton algorithms. In the case of nonlinear programming problems, assuming the gradients of active constraints to be linearly independent, the strong regularity reduces to some strong second-order sufficient condition.

The quadratic rate of convergence under the weak second-order sufficiency condition for nonlinear programming problems, and assuming the linear independence of the gradients of active constraints, has been recently obtained by the author [4]. This suggests that the theory of Newton-type methods for variational inequalities can be extended. For this purpose we use the new concept of semistability. We say that a solution $\bar{x}$ of a variational inequality is semistable if, given a small perturbation on the right-hand side, a solution x of the perturbed variational inequality that is sufficiently close to $\bar{x}$, is such that the distance of x to $\bar{x}$ is of the order of the magnitude of the perturbation. This does not imply the existence of a solution for the perturbed problem. Indeed, we give a counterexample showing that existence for a small perturbation does not always hold under the semistability hypothesis. We use a "hemistability" hypothesis in order to prove the existence of the sequence satisfying the Newton-type steps, then we show that semistability allows us to obtain in a simple way quadratic convergence for Newton's method and superlinear convergence for a large class of Newton-type algorithms (here we extend the Dennis and Moré [9] sufficient condition for superlinear convergence). This allows us to adapt Grzegòrski's [12] theory in order to derive the superlinear convergence of a large class of quasi-Newton updates including Broyden's one [7]. For polyhedral convex sets we may characterize semistability: it reduces to the condition that the solution $\bar{x}$ is an isolated solution of the variational inequality linearized at $\bar{x}$. An equivalent condition is the "strong positivity condition" of Reinoza [18]. We also check that for nondifferentiable data the theory can be extended using point-based approximations (reminiscent of those of Robinson [23]) that play the role of a linearized function.

The second part of this paper is devoted to a special class of variational inequalities generalizing optimality systems. The unknowns here are couples (x, λ) and we try to obtain conditions related to the superlinear convergence of $\{x^k\}$ alone. Indeed, we give a characterization of the superlinear convergence of $\{x^k\}$, valid under a second-order hypothesis satisfied by optimality systems for which the weak second-order sufficiency condition holds. This allows us to extend to inequality-constrained problems the characterization of Boggs *et al.* [3] for equality-constrained problems (this improves some previous results of the author [4] in which some necessary or sufficient conditions are given); our result assumes only that the gradients of active constraints are linearly independent and the weak second-order sufficient condition holds, but includes no strict complementarity hypothesis. We apply this result in order to obtain superlinear convergence for a large class of quasi-Newton updates. We note that these results can be used in order to formulate some globally convergent algorithms having fast convergence rates (see [5]).

2. Newton-Type Methods for Variational Inequalities

Let φ be a continuously differentiable mapping from $\mathbb{R}^q$ into $\mathbb{R}^q$. Given a closed convex subset K of $\mathbb{R}^q$ we consider the variational inequality

$$\langle \varphi(z), y - z \rangle \geq 0, \qquad \forall y \in K, \quad z \in K. \tag{2.1}$$

We may define the (closed convex) cone of outward normals to K at a point $z \in K$,

$$N(z) := \{x \in \mathbb{R}^q;\ \langle x, y - z \rangle \leq 0,\ \forall y \in K\},$$

and if $z \notin K$, $N(z) := \varnothing$. A relation equivalent to (2.1) is then

$$\varphi(z) + N(z) \ni 0. \tag{2.2}$$

When $K = \mathbb{R}^q$, $N(z) = \{0\}$ and we recover the equation $\varphi(z) = 0$. A natural extension of the Lagrange–Newton method for nonlinear programming (see [10]) is what we call the Newton-type algorithm:

Algorithm 1

0. Choose $z^0 \in \mathbb{R}^n$; $k \leftarrow 0$.
1. While z^k is not a solution of (2.2): choose M^k, a $q \times q$ matrix, and compute z^{k+1} solution of

$$\varphi(z^k) + M^k(z^{k+1} - z^k) + N(z^{k+1}) \ni 0. \tag{2.3}$$

We define Newton's method as Algorithm 1 when $M^k = \varphi'(z^k)$. In order to obtain estimates of the rate of convergence of $\{z^k\}$ we essentially use the following concept.

Definition 2.1. A solution $\bar{z}$ of (2.2) is said to be semistable if $c_1 > 0$ and $c_2 > 0$ exist such that, for all $(z, \delta) \in \mathbb{R}^q \times \mathbb{R}^q$, solution of

$$\varphi(z) + N(z) \ni \delta,$$

and $\|z - \bar{z}\| \leq c_1$, then $\|z - \bar{z}\| \leq c_2 \|\delta\|$.

Remark 2.1. (i) Note that this definition involves only those δ for which $\|\delta\| \le c_1/c_2$, because otherwise $\|z - \bar{z}\| \le c_2\|\delta\|$ is always satisfied whenever $\|z - \bar{z}\| \le c_1$; hence taking c_1 small enough, we can restrict δ to an arbitrary neighborhood of 0.

(ii) If $K = \mathbb{R}^q$ this condition reduces to the invertibility of $\varphi'(\bar{z})$; this is obtained as a consequence of Theorem 3.1.

Theorem 2.1. *Let $\bar{z}$ be a semistable solution of* (2.1), *and let $\{z^k\}$ computed by Algorithm* 1 *converge toward $\bar{z}$. Then*:

(i) *If $(\varphi'(\bar{z}) - M^k)(z^{k+1} - z^k) = o(z^{k+1} - z^k)$, then $\{z^k\}$ converges superlinearly.*
(ii) *If $(\varphi'(\bar{z}) - M^k)(z^{k+1} - z^k) = O(\|z^{k+1} - z^k\|^2)$ and φ' is locally Lipschitz, then $\{z^k\}$ converges quadratically.*

Proof. Define $\theta^k := (\varphi'(\bar{z}) - M^k)(z^{k+1} - z^k)$. We can write the Newton-type step (2.3) as

$$\varphi(z^{k+1}) + N(z^{k+1}) \ni r^k \tag{2.4}$$

with

$$\begin{aligned} r^k &:= \theta^k + \varphi(z^{k+1}) - \varphi(z^k) - \varphi'(\bar{z})(z^{k+1} - z^k) \\ &= \theta^k + o(z^{k+1} - z^k). \end{aligned}$$

If $\theta^k = o(z^{k+1} - z^k)$, then from the semistability of $\bar{z}$ and (2.4) we get

$$z^{k+1} - \bar{z} = O(r^k) = o(z^{k+1} - z^k) = o(\|z^{k+1} - \bar{z}\| + \|z^k - \bar{z}\|),$$

hence $z^{k+1} - \bar{z} = o(z^k - \bar{z})$, i.e., $\{z^k\}$ converges superlinearly. This proves (i). If φ' is locally Lipschitz and $\theta^k = O(\|z^{k+1} - z^k\|^2)$ we already know that $\{z^k\}$ converges superlinearly, hence $\|z^{k+1} - z^k\|/\|z^k - \bar{z}\| \to 1$. Let L be a Lipschitz constant of φ' at $\bar{z}$. We have, for k large enough,

$$\begin{aligned} &\|\varphi(z^{k+1}) - \varphi(z^k) - \varphi'(\bar{z})(z^{k+1} - z^k)\| \\ &\quad = \left\| \int_0^1 [\varphi'(z^k + \sigma(z^{k+1} - z^k)) - \varphi'(\bar{z})](z^{k+1} - z^k)\, d\sigma \right\| \\ &\quad \le L \max(\|z^{k+1} - \bar{z}\|, \|z^k - \bar{z}\|)\|z^{k+1} - z^k\| \\ &\quad \le 2L\|z^{k+1} - z^k\|^2, \end{aligned}$$

hence

$$z^{k+1} - \bar{z} = O(r^k) = O(\|z^{k+1} - z^k\|^2) = O(\|z^k - \bar{z}\|^2),$$

from which the quadratic convergence follows. □

Remark 2.2. Taking $K = \{0\}$ we see that the conditions of Theorem 2.1 are not necessary in general. However, when $K = \mathbb{R}^q$ (a case of a nonlinear equation) it is known that condition (i) is a characterization of superlinear convergence [8].

Corollary 2.1. *If $\{z^k\}$ computed by Algorithm 1 converges toward a semistable solution $\bar{z}$ of (2.1), then:*

(i) *If $M^k \to \varphi'(\bar{z})$, $\{z^k\}$ converges superlinearly.*
(ii) *If φ' is locally Lipschitz and $M^k = \varphi'(\bar{z}) + O(z^k - \bar{z})$ (which is the case for Newton's method under the hypothesis of Lipschitz continuity of φ'), then $\{z^k\}$ converges quadratically.*

Until now we assumed the existence of a converging sequence instead of giving the hypotheses that imply its existence. Our point of view is that it is clearer to do so; indeed, if we now want to prove that the sequence is well defined for, say, Newton's method with a good starting point, we just have to posit the following definition:

Definition 2.2. We say that $\bar{z}$ is a hemistable solution of (2.1) if, for all $\alpha > 0$, $\varepsilon > 0$ exists such that, given $\hat{z} \in \mathbb{R}^q$, the variational inequality (in z)

$$\varphi(\hat{z}) + M(z - \hat{z}) + N(z) \ni 0$$

has a solution z satisfying $\|z - \bar{z}\| \le \alpha$, whenever $\|\hat{z} - \bar{z}\| + \|M - \varphi'(\bar{z})\| < \varepsilon$.

Then, using Corollary 2.1, we obtain

Theorem 2.2 (Local Analysis of Newton's Method). *If $\bar{z}$ is a semistable and hemistable solution of (2.1), $\varepsilon > 0$ exists such that if $\|z^0 - \bar{z}\| \le \varepsilon$, then:*

(i) *At each step k a z^{k+1} solution of the Newton step satisfying $\|z^{k+1} - z^k\| \le 2\varepsilon$ exists.*
(ii) *The sequence $\{z^k\}$ defined in this way converges superlinearly (quadratically if φ' is locally Lipschitz) toward $\bar{z}$.*

Proof. We just have to prove (i) and the convergence of $\{z^k\}$ toward $\bar{z}$, then (ii) will follow from Corollary 2.1. Assume φ' is merely continuous at $\bar{z}$. Take $\varepsilon_0 \le \min(c_1, 1/3c_2)$ where c_1, c_2 are given by the semistability condition. From the hemistability condition we have that, for some $\varepsilon \in (0, c_1)$, $\|z^k - \bar{z}\| \le \varepsilon$ implies the existence of z^{k+1} such that $\|z^{k+1} - \bar{z}\| \le \varepsilon_0$ and

$$\varphi(z^k) + \varphi'(z^k)(z^{k+1} - z^k) + N(z^{k+1}) \ni 0. \tag{2.5}$$

Now $\varphi(z^{k+1}) + N(z^{k+1}) \ni \delta^k$ where

$$\delta^k := \varphi(z^{k+1}) - \varphi(z^k) - \varphi'(z^k)(z^{k+1} - z^k). \tag{2.6}$$

From differential calculus we obtain, reducing ε_0 and ε if necessary, that

$$\|\delta^k\| \le \frac{1}{3c_2} \|z^{k+1} - z^k\|. \tag{2.7}$$

As $\varepsilon_0 \le c_1$ the semistability condition gives

$$\|z^{k+1} - \bar{z}\| \le \tfrac{1}{3}\|z^{k+1} - z^k\| \le \tfrac{1}{3}\|z^{k+1} - \bar{z}\| + \tfrac{1}{3}\|z^k - \bar{z}\|,$$

hence

$$\|z^{k+1} - \bar{z}\| \le \tfrac{1}{2}\|z^k - \bar{z}\| \tag{2.8}$$

and

$$\|z^{k+1} - z^k\| \le \|z^{k+1} - \bar{z}\| + \|z^k - \bar{z}\| \le 2\varepsilon.$$

This proves (i) and the linear convergence of $\{z^k\}$. □

Remark 2.3. The condition $\|z^{k+1} - z^k\| \le 2\varepsilon$ in Theorem 2.2 is constructive in the sense that if we choose the solution of (2.5) closest to z^k, then, if the starting point z^0 is close enough to $\bar{z}$, the condition is satisfied and the conclusion of Theorem 2.2 follows.

Remark 2.4. (i) Semistability does not imply hemistability, as is shown by the following example. Consider the variational inequality with $K = \mathbb{R}^+$:

$$-z + N(z) \ni 0,$$

corresponding to the optimality system of the ill-posed optimization problem

$$\min\left\{\frac{-z^2}{2};\ z \ge 0\right\}.$$

Here

$$N(z) = \begin{cases} \varnothing & \text{if } z < 0, \\ \mathbb{R}^+ & \text{if } z = 0, \\ 0 & \text{if } z > 0. \end{cases}$$

We have that $\bar{z} = 0$ is the unique solution. Now the perturbed variational inequality

$$-z + N(z) \ni \delta$$

has a solution iff $\delta \le 0$ and this solution is $z = -\delta$, hence semistability holds although the variational inequality may have no solution for $\|\delta\|$ arbitrarily small.

Let us now prove that hemistability does not hold. Here $\varphi(z) = -z$ and $\varphi'(\bar{z}) = -1$; take $\hat{z} = \varepsilon$ and $M = \varepsilon - 1$ with $\varepsilon \in (0, 1)$; we discuss the solvability near 0 of

$$-\varepsilon + (\varepsilon - 1)(z - \varepsilon) + \partial\mathbb{R}^+(z) \ni 0.$$

If z is a solution, either $z = 0$, but then $-\varepsilon + (\varepsilon - 1)(z - \varepsilon) = -\varepsilon^2 < 0$, impossible; or $-\varepsilon + (\varepsilon - 1)(z - \varepsilon) = 0$, i.e., $z = \varepsilon^2/(\varepsilon - 1) < 0$, which is also impossible. Hence the perturbed variational inequality has no solution, although $(\hat{z}, M)$ is arbitrarily close to $(\bar{z}, \varphi(\bar{z}))$.

(ii) A sufficient condition for semistability and hemistability is the strong regularity of Robinson [19]. Indeed, strong regularity amounts to saying that the equation

$$\varphi(\bar{z}) + \varphi'(\bar{z})(z - \bar{z}) + N(z) \ni \delta$$

is such that $\varepsilon > 0$, $\alpha > 0$, $\beta > 0$ exist such that if $\|\delta\| \le \varepsilon$, then a unique solution z exists such that $\|z - \bar{z}\| \le \alpha$, and this z satisfies $\|z - \bar{z}\| \le \beta\|\delta\|$. Now let z solve the perturbed variational inequality $\varphi(z) + N(z) \ni \delta$. Then

$$\varphi(\bar{z}) + \varphi'(\bar{z})(z - \bar{z}) + N(z) \ni \delta + o(z - \bar{z}).$$

Strong regularity implies that $z - \bar{z} = O(\delta) + o(z - \bar{z})$, hence $z - \bar{z} = O(\delta)$, i.e., the semistability holds.

Also if $\bar{z}$ is a strongly regular solution of (2.2) it is obviously a strongly regular solution of the linearized variational inequality

$$\varphi(\bar{z}) + \varphi'(\bar{z})(z - \bar{z}) + N(z) \ni 0.$$

We apply Theorem 2.1 of [19]. If $\|\hat{z} - \bar{z}\| + \|M - \varphi'(\bar{z})\|$ is small enough, the variational inequality

$$\varphi(\hat{z}) + M(z - \hat{z}) + N(z) \ni 0$$

has a solution and

$$\|z - \bar{z}\| = O(\varphi(\hat{z}) - \varphi(\bar{z}));$$

this implies hemistability.

(iii) We see later that in the case of optimality systems for local solutions of nonlinear programming problems, semistability and hemistability are equivalent.

Theorem 2.1 may also be used in order to derive superlinear convergence of some quasi-Newton algorithm. By quasi-Newton algorithm we mean a Newton-type algorithm with M^{k+1} satisfying the so-called quasi-Newton equation

$$M(z^{k+1} - z^k) = \varphi(z^{k+1}) - \varphi(z^k). \tag{2.9}$$

A typical situation is when a closed convex subset $\mathscr{K}$ of the space of $q \times q$ matrices is known to satisfy

$$\varphi'(z) \in \mathscr{K}, \qquad \forall z \in \mathbb{R}^q. \tag{2.10}$$

Then M^{k+1} is taken as a solution of

$$\min\|M - M^k\|_{\#}; \qquad M \in \mathscr{K} \quad \text{and} \quad M \text{ satisfies (2.9)}. \tag{2.11}$$

Here $\|\cdot\|_{\#}$ is a matrix norm that we assume to be associated with a scalar product. If $\|\cdot\|_{\#}$ is the Frobenius norm we recover Broyden's update when $\mathscr{K}$ is the space of $q \times q$ matrices, the PSB update when $\mathscr{K}$ is the space of symmetric matrices, etc.; see [12]. We first quote

Lemma 2.1. *Under the hypotheses of Theorem* 2.1, *if* $\{M^k\}$ *satisfies the quasi-Newton equation and*

$$(M^{k+1} - M^k)(z^{k+1} - z^k) = o(z^{k+1} - z^k),$$

then $\{z^k\}$ *converges superlinearly.*

Proof. Using (2.9) we get

$$\begin{aligned}(M^{k+1} - M^k)(z^{k+1} - z^k) &= \varphi(z^{k+1}) - \varphi(z^k) - M^k(z^{k+1} - z^k) \\ &= (\varphi'(\bar{z}) - M^k)(z^{k+1} - z^k) + o(z^{k+1} - z^k).\end{aligned}$$

The conclusion is then obtained with Theorem 2.1. □

Theorem 2.3. *Let* φ' *be locally Lipschitz, and let* $\bar{z}$ *be a semistable and hemistable solution of* (2.2). *We assume that* (2.9)–(2.11) *hold. Then* $\varepsilon > 0$ *exists such that if*

$$\|z^0 - \bar{z}\| + \|M^0 - \varphi'(\bar{z})\|_\# < \varepsilon,$$

then:

(i) *At each step* k *a* z^{k+1} *solution of the Newton-type step satisfying* $\|z^{k+1} - z^k\| \le 2\varepsilon$ *exists.*
(ii) *The sequence* $\{z^k\}$ *defined in this way converges superlinearly toward* $\bar{z}$.

Proof. Define

$$S^k := \{M \in \mathscr{K};\ M(z^{k+1} - z^k) = \varphi(z^{k+1}) - \varphi(z^k)\}.$$

Then M^{k+1} is the projection of M^k onto S^k (with the $\|\cdot\|_\#$ norm), hence for all $M \in S^k$ we have (see Theorem 1 of [12])

$$\|M^{k+1} - M^k\|_\#^2 + \|M^{k+1} - M\|_\#^2 \le \|M^k - M\|_\#^2 \tag{2.12}$$

and *a fortiori*

$$\|M^{k+1} - M\|_\# \le \|M^k - M\|_\#. \tag{2.13}$$

Define

$$\psi^k := \int_0^1 \varphi'(z^k + \sigma(z^{k+1} - z^k))\, d\sigma, \tag{2.14}$$

$$v^k := \max(\|z^{k+1} - \bar{z}\|, \|z^k - \bar{z}\|). \tag{2.15}$$

Then ψ^k is an element of S^k and, for k large enough, we have, L being a Lipschitz constant of φ' in a neighborhood of $\bar{z}$ in the $\|\cdot\|_\#$ norm,

$$\|\psi^k - \varphi'(\bar{z})\|_\# = \left\| \int_0^1 [\varphi'(z^k + \sigma(z^{k+1} - z^k)) - \varphi'(\bar{z})]\, d\sigma \right\|_\# \le Lv^k,$$

hence, taking $M = \psi^k$ in (2.13), and using the previous inequality, we get

$$\|M^{k+1} - \varphi'(\bar{z})\|_{\#} \leq \|M^k - \varphi'(\bar{z})\|_{\#} + 2Lv^k. \tag{2.16}$$

We prove in Lemma 2.2 below that this bounded deterioration result implies that (for ε small enough) $z^k \to \bar{z}$ linearly and $\|M^k - \varphi'(\bar{z})\|_{\#}$ converges. As $\psi^k \to \varphi'(\bar{z})$, $\|M^k - \psi^k\|_{\#}$ and $\|M^{k+1} - \psi^k\|_{\#}$ also converge toward the same limit. Taking $M = \psi^k$ in (2.12) we deduce that $\|M^{k+1} - M^k\| \to 0$; this and Lemma 2.1 imply the conclusion. □

Lemma 2.2 (Linear Convergence under Bounded Deterioration). *Let $\bar{z}$ be as in Theorem* 2.3. *Let $\{z^k\}$ be computed by a Newton-type algorithm such that $\{M^k\}$ satisfies* (2.16). *Then, for any θ in* (0, 1), *$\varepsilon > 0$ exists such that if*

$$\|z^0 - \bar{z}\| + \|M^0 - \varphi'(\bar{z})\|_{\#} < \varepsilon,$$

then:

(i) *At each step k a z^{k+1} solution of the Newton-type step satisfying $\|z^{k+1} - z^k\| \leq 2\varepsilon$ exists.*
(ii) *$z^k \to \bar{z}$ linearly with speed θ, i.e., $\|z^{k+1} - \bar{z}\| \leq \theta\|z^k - \bar{z}\|$.*
(iii) *$\|M^k - \varphi'(\bar{z})\|_{\#}$ converges.*

Proof. Writing (2.3) as

$$\varphi(z^k) + \varphi'(\bar{z})(z^{k+1} - z^k) + N(z^{k+1}) \ni (\varphi'(\bar{z}) - M^k)(z^{k+1} - z^k),$$

and using

$$\varphi(z^{k+1}) = \varphi(z^k) + \varphi'(\bar{z})(z^{k+1} - z^k) + \int_0^1 [\varphi'(z^k + \sigma(z^{k+1} - z^k)) - \varphi'(\bar{z})](z^{k+1} - z^k)\, d\sigma,$$

we deduce that

$$\varphi(z^{k+1}) + N(z^{k+1}) \ni \delta^k$$

with (v^k being defined in (2.15) and using the canonical norm of $L(\mathbb{R}^n)$)

$$\|\delta^k\| \leq (\|\varphi'(\bar{z}) - M^k\| + Lv^k)\|z^{k+1} - z^k\|,$$

and from the semistability hypothesis we deduce

$$\|z^{k+1} - \bar{z}\| \leq c_2(\|\varphi'(\bar{z}) - M^k\| + Lv^k)\|z^{k+1} - z^k\|.$$

Using the triangle inequality

$$\|z^{k+1} - z^k\| \leq \|z^{k+1} - \bar{z}\| + \|z^k - \bar{z}\|$$

we deduce that whenever

$$\|\varphi'(\bar{z}) - M^k\| + Lv^k < \frac{1}{c_2},$$

then

$$\|z^{k+1} - \bar{z}\| \le \theta_1 \|z^k - \bar{z}\|$$

with

$$\theta_1 = \frac{c_2(\|\varphi'(\bar{z}) - M^k\| + Lv^k)}{1 - c_2(\|\varphi'(\bar{z}) - M^k\| + Lv^k)}.$$

Using the hemistability hypothesis in order to estimate v^k we see that $\varepsilon_0 > 0$ exists such that $\theta_1 \le \theta$ whenever

$$\|\varphi'(\bar{z}) - M^k\|_{\#} + \|z^{k+1} - \bar{z}\| \le \varepsilon_0. \tag{2.17}$$

If $\varepsilon \le \varepsilon_0$ this is the case for $k = 0$. Now assume that (2.17) is satisfied for $k = 0, \dots, \bar{k}$. Then with (2.16) and using the linear convergence of $\{z^k\}$, we get

$$\|z^k - \bar{z}\| \le \theta^k \varepsilon, \qquad k = 0 \quad \text{to} \quad \bar{k} + 1,$$

$$v^k \le 2 \sum_{k=0}^{\bar{k}+1} \|z^k - \bar{z}\| \le 2 \sum_{k=0}^{\infty} \theta^k \|z^0 - \bar{z}\| \le \frac{2\varepsilon}{1 - \theta},$$

$$\begin{aligned} \|\varphi'(\bar{z}) - M^{\bar{k}+1}\|_{\#} &\le \|\varphi'(\bar{z}) - M^0\|_{\#} + 2L \sum_{k=0}^{\bar{k}+1} v^k \\ &\le \varepsilon + \frac{4L\varepsilon}{1 - \theta}, \end{aligned}$$

hence

$$\|\varphi'(\bar{z}) - M^{\bar{k}+1}\|_{\#} + \|z^{\bar{k}+1} - \bar{z}\| \le 2 \frac{\varepsilon + 4L\varepsilon}{1 - \theta} \le \frac{4L + 2}{1 - \theta} \varepsilon. \tag{2.18}$$

We now choose

$$\varepsilon = \frac{1 - \theta}{4L + 2} \varepsilon_0.$$

For this value it appears that (2.17) is also satisfied for $k = \bar{k} + 1$, hence (by recurrence) for all $k \in \mathbb{N}$. This proves the linear convergence with speed θ. Also, for all $k \in \mathbb{N}$ and $l < k$,

$$\begin{aligned} \|\varphi'(\bar{z}) - M^k\|_{\#} &\le \|\varphi'(\bar{z}) - M^l\|_{\#} + 2L \sum_{i=l}^{k-1} v^i \\ &\le \|\varphi'(\bar{z}) - M^l\|_{\#} + \frac{2L\theta^l}{1 - \theta}, \end{aligned}$$

hence

$$\overline{\lim}\|\varphi'(\bar{z}) - M^k\|_{\#} \le \|\varphi'(\bar{z}) - M^l\|_{\#} + \frac{2L\theta^l}{1-\theta}.$$

When $l \to \infty$ we deduce

$$\overline{\lim}\|\varphi'(\bar{z}) - M^k\|_{\#} \le \underline{\lim}\|\varphi'(\bar{z}) - M^k\|_{\#},$$

i.e., $\|\varphi'(\bar{z}) - M^k\|_{\#}$ converges. □

Remark 2.5. The hemistability hypothesis is needed only to ensure the existence of z^{k+1} close to $\bar{z}$. The rest of the analysis relies upon the semistability hypothesis.

3. Characterization of Semistability when *K* is Polyhedral

We assume here that K is polyhedral, i.e., defined by a finite number of linear equalities and inequalities. This allows us to give several characterizations of semistability.

Theorem 3.1. *If K is polyhedral and $\bar{z}$ is a solution of* (2.1), *$\bar{z}$ is semistable iff one of the following hypotheses holds*:

(a) *$\bar{z}$ is an isolated solution of the linearization at $\bar{z}$ of* (2.2):

$$\varphi(\bar{z}) + \varphi'(\bar{z})(z - \bar{z}) + N(z) \ni 0. \tag{3.1}$$

(b) *We have $\langle z - \bar{z}, \varphi'(\bar{z})(z - \bar{z})\rangle > 0$ for all $z \in K$ different to the $\bar{z}$ solution of*

$$\langle \varphi(\bar{z}), z - \bar{z}\rangle = 0, \tag{3.2i}$$

$$\varphi(\bar{z}) + \varphi'(\bar{z})(z - \bar{z}) + N(\bar{z}) \ni 0. \tag{3.2ii}$$

(c) *The conditions below have no solution other than $\bar{z}$*:

$$N(z) \subset N(\bar{z}), \tag{3.3i}$$

$$\langle \varphi(\bar{z}), z - \bar{z}\rangle = 0, \tag{3.3ii}$$

$$\alpha\varphi(\bar{z}) + \varphi'(\bar{z})(z - \bar{z}) + N(z) \ni 0 \quad \textit{for some} \quad \alpha \ge 0. \tag{3.3iii}$$

Remark 3.1. (i) In the case of a nonlinear equation it follows from condition (a) that semistability is equivalent to the inversibility of the Jacobian, which in turn is also equivalent to hemistability.

(ii) Reinoza [18, Theorem 2.1] has already proved the equivalence of (a) and (b). He called condition (b) a strong positivity condition, although in the context of nonlinear programming we will see that it corresponds to weak second-order sufficient conditions; hence it might be better to call it a weak positivity condition.

Proof of Theorem 3.1. We prove that

$$\{\bar{z} \text{ is semistable}\} \Rightarrow \text{(a)} \Rightarrow \text{(b)} \Rightarrow \text{(c)} \Rightarrow \{\bar{z} \text{ is semistable}\}.$$

(a) Proof of $\{\bar{z}$ semistable$\} \Rightarrow$ (a). If z is solution of (3.1), then, from the first-order expansion of φ at $\bar{z}$,

$$\varphi(z) + N(z) \ni o(z - \bar{z}),$$

hence if $\bar{z}$ is semistable and $\|z - \bar{z}\| \leq c_1$, we get $\|z - \bar{z}\| = o(z - \bar{z})$ and this implies $z = \bar{z}$ for z close enough to $\bar{z}$; hence (a) holds.

(b) Proof of (a) $\Rightarrow$ (b). Let z in K contradict (b), i.e., $z \neq \bar{z}$, z satisfies (3.2) but $\langle z - \bar{z}, \varphi'(\bar{z})(z - \bar{z})\rangle \leq 0$. From (3.2) we get

$$0 \leq \langle \varphi(\bar{z}) + \varphi'(\bar{z})(z - \bar{z}), z - \bar{z}\rangle = \langle z - \bar{z}, \varphi'(\bar{z})(z - \bar{z})\rangle,$$

hence

$$\langle z - \bar{z}, \varphi'(\bar{z})(z - \bar{z})\rangle = 0. \tag{3.4}$$

For α in $]0, 1[$ define $z^\alpha := \bar{z} + \alpha(z - \bar{z})$. From (3.2ii), (2.2), and the convexity of $N(\bar{z})$ we deduce that

$$\varphi(\bar{z}) + \varphi'(\bar{z})(z^\alpha - \bar{z}) + N(\bar{z}) \ni 0,$$

hence with (3.2i) and (3.4), for all $y \in K$,

$$\begin{aligned} 0 &\leq \langle \varphi(\bar{z}) + \varphi'(\bar{z})(z^\alpha - \bar{z}), y - \bar{z}\rangle \\ &= \langle \varphi(\bar{z}) + \varphi'(\bar{z})(z^\alpha - \bar{z}), y - z^\alpha\rangle, \end{aligned}$$

that is,

$$\varphi(\bar{z}) + \varphi'(\bar{z})(z^\alpha - \bar{z}) + N(z^\alpha) \ni 0,$$

hence z^α is a solution of (3.1). Also $z^\alpha \to \bar{z}$ when $\alpha \searrow 0$; this contradicts (a).

(c) Proof of (b) $\Rightarrow$ (c). Assume that (c) does not hold and let $z \in K$, $z \neq \bar{z}$ be a solution of (3.3). From (3.3ii) and (3.3iii) we deduce that

$$\langle z - \bar{z}, \varphi'(\bar{z})(z - \bar{z})\rangle \leq 0.$$

As (3.2i) coincides with (3.3ii) it remains to derive (3.2ii) in order to get a contradiction with (b). If $\alpha \leq 1$, multiplying relation (2.2) by $(1 - \alpha)$, adding it to (3.3iii), and using (3.3i) we get (3.2ii). If $\alpha > 1$ we may check similarly, dividing (3.3iii) by α, that $y^\alpha := \bar{z} + (1/\alpha)(z - \bar{z})$ contradicts (b).

(d) Proof of (c) $\Rightarrow$ $\{\bar{z}$ is semistable$\}$. If $\bar{z}$ is not semistable let $z^k \to \bar{z}$ and $\delta^k \to 0$ in $\mathbb{R}^n$ be such that

$$\varphi(z^k) + N(z^k) \ni \delta^k, \tag{3.5}$$

and $\|\delta^k\|/\|z^k - \bar{z}\| \to 0$. Define $\beta^k := \|z^k - \bar{z}\|^{-1}$ and $w^k := \beta^k(z^k - \bar{z})$. Then substituting $\varphi(\bar{z}) + \varphi'(\bar{z})(z^k - \bar{z}) + o(z^k - \bar{z})$ to $\varphi(z^k)$ in (3.5) we get, after multiplication by β^k,

$$\beta^k\varphi(\bar{z}) + \varphi'(\bar{z})w^k + N(z^k) \ni \beta^k\delta^k + \beta^k o(z^k - \bar{z}). \tag{3.6}$$

The right-hand side of (3.6) has limit 0. As K is a polyhedron we may extract, without loss of generality, a subsequence such that $N(z^0) = N(z^k)$ for all k; also $\|w^k\| = 1$, hence $\{w^k\}$ has at least a limit-point w (for the same subsequence) with $\|w\| = 1$. Again as K is a polyhedron, the set $N^0 := N(z^0) + \mathbb{R}^+\varphi(\bar{z})$ is the cone of exterior normals at z^0 to the set

$$K^0 := K \cap \{z \in \mathbb{R}^q;\ \langle z - z^0, \varphi(\bar{z})\rangle \leq 0\}.$$

Hence N^0 is closed. By (3.6) and the closedness of N^0 we have

$$\mathbb{R}^+\varphi(\bar{z}) + \varphi'(\bar{z})w + N(z^0) \ni 0. \tag{3.7}$$

Also as $\beta^k \geq 0$ and the vectors $\bar{z} + (\beta^k)^{-1}w^k = z^k$, $z^k - (\beta^k)^{-1}w^k = \bar{z}$ are elements of K, we get, from (2.1) and (3.5),

$$\begin{cases} \langle w^k, \varphi(\bar{z})\rangle = \beta^k\langle z^k - \bar{z}, \varphi(\bar{z})\rangle \geq 0, \\ -\langle w^k, \varphi(z^k)\rangle = \beta^k\langle \bar{z} - z^k, \varphi(z^k)\rangle \geq \beta^k\langle \bar{z} - z^k, \delta^k\rangle \to 0. \end{cases} \tag{3.8}$$

As $z^k \to \bar{z}$, $\varphi(z^k) \to \varphi(\bar{z})$. This, (3.8), and $w^k \to w$ imply

$$\langle w, \varphi(\bar{z})\rangle = 0. \tag{3.9}$$

Now, as K is a polyhedron, $\bar{z} + \varepsilon w$ is in K for $\varepsilon > 0$ small enough. Let us check that $N(\bar{z} + \varepsilon w) \supset N(z^0)$. It is sufficient to check that any linear inequality constraint defining K that is active at z^0 is also active at $\bar{z} + \varepsilon w$. Here we say that a constraint $\langle a, z\rangle \leq b$ is active at z if $\langle a, z\rangle = b$. Extracting again if necessary a subsequence we may assume that the set of active constraints is the same for all $\{z^k\}$. Then for the subsequence considered here we have $\langle a, z^k\rangle = b$, hence $\langle a, \bar{z}\rangle = b$ and $\langle a, w^k\rangle = 0$, from which $\langle a, w\rangle = 0$, and finally $\langle a, \bar{z} + \varepsilon w\rangle = b$. This proves that $N(\bar{z} + \varepsilon w) \supset N(z^0)$. This and (3.7) (multiplied by $\varepsilon > 0$) imply

$$\mathbb{R}^+\varphi(\bar{z}) + \varepsilon\varphi'(\bar{z})w + N(\bar{z} + \varepsilon w) \ni 0. \tag{3.10}$$

Also, for $\varepsilon > 0$ small enough and as K is a polyhedron, $N(\bar{z} + \varepsilon w) \subset N(\bar{z})$. This, (3.9), (3.10), and the fact that $z = \bar{z} + \varepsilon w$ is in K give a contradiction to (c). □

Remark 3.2. The proof of

$$\{\bar{z} \text{ is semistable}\} \Rightarrow \text{(a)} \Rightarrow \text{(b)} \Rightarrow \text{(c)}$$

does not use the fact that K is polyhedral.

4. Extension of the Theory to Nonsmooth Data

Although we are mainly interested in this paper by finite-dimensional variational inequalities with smooth data we give here a partial extension of the previous results to problems in a Hilbert space with nonsmooth data. Let K be a closed convex subset of a Hilbert space Z, let $N(z)$ be the cone of outward normals to K at z, and let φ be a mapping from Z into itself. In order to define an extension of Algorithm 1 for the problem

$$\varphi(z) + N(z) \ni 0, \tag{4.1}$$

we use a concept of point-based approximation (PBA) close to the one of Robinson [23].

Definition 4.1. We say that $\psi: Z \times Z \to Z$ is a PBA to φ if, for any two sequences $\{y^k\}$, $\{z^k\}$ converging to the same point, the following holds:

$$\|\varphi(y^k) - \psi(z^k, y^k)\| \le r(y^k, z^k), \tag{4.2}$$

with $r(y^k, z^k)/\|y^k - z^k\| \to 0$.

Here $\psi(z^k, \cdot)$ can be seen as a generalization of the linearization of φ at z^k (see Remark 4.1 below). We now define a somewhat abstract Newton-type method as the following algorithm:

Algorithm 2

0. Choose $z^0 \in Z$; $k \leftarrow 0$.
1. While z^k does not satisfy (4.1): choose a mapping $\Xi^k: Z \to Z$, an approximation of $\psi(z^k, \cdot)$. Compute the z^{k+1} solution of

$$\Xi^k(z^{k+1}) + N(z^{k+1}) \ni 0. \tag{4.3}$$

We define semistability as in Section 2.

Theorem 4.1. *If $\{z^k\}$ computed by Algorithm 2 converges toward a semistable solution $\bar{z}$ of (4.1), then:*

(i) *If $\psi(z^k, z^{k+1}) - \Xi^k(z^{k+1}) = o(z^{k+1} - z^k)$, then $\{z^k\}$ converges superlinearly.*
(ii) *If $\psi(z^k, z^{k+1}) - \Xi^k(z^{k+1}) = O(\|z^{k+1} - z^k\|^2)$ and, for some $c_1 > 0$ and all (y, z) close enough to $\bar{z}$, the function r in (4.2) satisfies $r(y, z) \le c_1\|y - z\|^2$, then $\{x^k\}$ converges quadratically.*

Proof. Writing step (4.3) as

$$\psi(z^k, z^{k+1}) + N(z^{k+1}) \ni \psi(z^k, z^{k+1}) - \Xi^k(z^{k+1})$$

and using (4.2), we deduce that

$$\varphi(z^{k+1}) + N(z^{k+1}) \ni \psi(z^k, z^{k+1}) - \Xi^k(z^{k+1}) + o(z^{k+1} - z^k).$$

In case (i) it follows from semistability that $z^{k+1} - \bar{z} = o(z^{k+1} - z^k)$, hence z^k converges superlinearly. In case (ii) we similarly obtain $z^{k+1} - \bar{z} = O(\|z^{k+1} - z^k\|^2)$, which implies the quadratic convergence. □

Remark 4.1. Theorem 4.1 can be seen as an extension of Theorem 2.1. Indeed, if φ is continuously differentiable and

$$\psi(z^k, z^{k+1}) = \varphi(z^k) + \varphi'(z^k)(z^{k+1} - z^k),$$

$$\Xi^k(z^{k+1}) = \varphi(z^k) + M^k(z^{k+1} - z^k)$$

for some $q \times q$ matrix M^k, then

$$\begin{aligned}\psi(z^k, z^{k+1}) - \Xi^k(z^{k+1}) &= (\varphi'(z^k) - M^k)(z^{k+1} - z^k)\\ &= (\varphi'(\bar{z}) - M^k)(z^{k+1} - z^k) + o(z^{k+1} - z^k),\end{aligned}$$

hence point (i) of Theorem 4.1 reduces to point (i) of Theorem 2.1. Similarly, if φ' is locally Lipschitz we have

$$(\varphi(z^k) - M^k)(z^{k+1} - z^k) = (\varphi'(\bar{z}) - M^k)(z^{k+1} - z^k) + O(\|z^k - \bar{z}\| \cdot \|z^{k+1} - z^k\|),$$

the last term being $O(\|z^{k+1} - z^k\|)^2$ as $\|z^{k+1} - z^k\|/\|z^k - \bar{z}\| \to 1$ because of the superlinear convergence, hence point (ii) of Theorem 4.1 reduces to point (ii) of Theorem 2.1.

We define the directional derivatives $\varphi'(\cdot, \cdot)$ of φ as the limit

$$\varphi'(z, d) := \lim_{\alpha \searrow 0} \frac{1}{\alpha} [\varphi(z + \alpha d) - \varphi(z)].$$

We state in Theorem 4.2 below an extension of Theorem 3.1. Theorem 4.2 applies to B-differentiable mappings (here B stands for Bouligand), as defined by Robinson [22], i.e., mappings having the following property: φ is locally Lipschitz, has directional derivatives, and $d \to \varphi'(x, d)$ is Lipschitz. Then it is known that (for given x) $\varphi(x + d) = \varphi(x) + \varphi'(x, d) + o(d)$ (see also [24]).

Theorem 4.2. *Assume that $Z = \mathbb{R}^q$, φ is a B-differentiable mapping, K is polyhedral, and $\bar{z}$ is a solution of* (4.1). *Then $\bar{z}$ is semistable iff one of the following hypotheses holds*:

(a) *$\bar{z}$ is an isolated solution of the linearization at $\bar{z}$ of* (4.1) *defined as follows*:

$$\varphi(\bar{z}) + \varphi'(\bar{z}, z - \bar{z}) + N(z) \ni 0.$$

(b) *We have $\langle z - \bar{z}, \varphi'(\bar{z}, z - \bar{z}) \rangle > 0$ for all z different to the $\bar{z}$ solution of*

$$\langle \varphi(\bar{z}), z - \bar{z} \rangle = 0,$$

$$\varphi(\bar{z}) + \varphi'(\bar{z}, z - \bar{z}) + N(\bar{z}) \ni 0.$$

(c) *The relation below has no solution other than* $\bar{z}$:

$$N(z) \subset N(\bar{z}),$$

$$\langle \varphi(\bar{z}), z - \bar{z} \rangle = 0,$$

$$\alpha\varphi(\bar{z}) + \varphi'(\bar{z}, z - \bar{z}) + N(z) \ni 0 \qquad \textit{for some} \quad \alpha \geq 0.$$

The proof is the same as that of Theorem 3.1, replacing first-order variations by directional derivatives.

5. Convergence Analysis for Some Structured Variational Inequalities

We now specialize our study to a particular case of variational inequalities. In the next section we apply the results of this section to nonlinear programming problems. Let F, g be smooth (resp. C^1 and C^2) mappings: $\mathbb{R}^n \to \mathbb{R}^n$ and $\mathbb{R}^n \to \mathbb{R}^p$, respectively. Let I, J be a partition of $\{1, \ldots, p\}$. By $g(x) \ll 0$ we mean

$$g_i(x) \leq 0, \qquad \forall i \in I,$$

$$g_j(x) = 0, \qquad \forall j \in J.$$

We now consider the system (in which $\lambda \in \mathbb{R}^p$)

$$\begin{cases} F(x) + g'(x)^*\lambda = 0, \\ g(x) \ll 0, \qquad \lambda_I \geq 0, \qquad \lambda_i g_i(x) = 0, \quad \forall i \in I. \end{cases} \tag{5.1}$$

As observed in [23] we may embed (5.1) into (2.1) in the following way. Put $q := n + p$, $z := (x, \lambda)$, and

$$\varphi(x, \lambda) := \begin{pmatrix} F(x) + g'(x)^*\lambda \\ -g(x) \end{pmatrix},$$

$$K_1 := \{\lambda \in \mathbb{R}^p, \lambda_I \geq 0\}, \qquad K := \mathbb{R}^n \times K_1,$$

so that K is polyhedral and

$$N(x, \lambda) = \{0\} \times N_1(\lambda),$$

with $N_1(\lambda)$ the normal cone (or cone of outwards normals) to K_1 at λ, i.e.,

$$N_1(\lambda) = \begin{cases} \varnothing \text{ if } \lambda \text{ is not in } K_1, \text{ otherwise} \\ \{\mu \in \mathbb{R}^p; \mu_J = 0; \mu_I \leq 0; \mu_i = 0 \text{ if } \lambda_i > 0, \forall i \in I\}. \end{cases}$$

The corresponding variational inequality can be written in the following way:

$$\begin{cases} F(x) + g'(x)^*\lambda = 0, \\ -g(x) + N_1(\lambda) \ni 0. \end{cases} \tag{5.2}$$

Let us denote

$$H(x, \lambda) := F'(x) + \sum_{i=1}^{p} \lambda_i \nabla^2 g_i(x).$$

Then we have

$$\varphi'(x, \lambda) = \begin{pmatrix} H(x, \lambda) & g'(x)^* \\ -g'(x) & 0 \end{pmatrix} \tag{5.3}$$

and

$$\langle (y, \mu), \varphi'(x, \lambda)(y, \mu) \rangle = \langle y, H(x, \lambda) y \rangle. \tag{5.4}$$

Taking (5.2), (5.3), and Theorem 3.1 into account, we see that the semistability for (5.2) (expressed at some point of the $(\bar{x}, \bar{\lambda})$ solution of (5.2)) can be stated as

$$\begin{cases} (y, \mu) = 0 \quad \text{is an isolated solution of} \\ H(\bar{x}, \bar{\lambda}) y + g'(\bar{x})^* \mu = 0, & (5.5\text{i}) \\ g(\bar{x}) + g'(\bar{x}) y \in N_1(\bar{\lambda} + \mu). & (5.5\text{ii}) \end{cases}$$

For any $\hat{I} \subset I$ by $z \overset{\hat{I}}{\leqslant} 0$ we mean $z_J = 0$ and $z_i \leq 0$ for all i in $\hat{I}$. Let us define

$$\bar{I} := \{ i \in I;\ g_i(\bar{x}) = 0 \},$$

$$I^+ := \{ i \in \bar{I};\ \bar{\lambda}_i > 0 \},$$

$$I^0 := \bar{I} - I^+ = \{ i \in \bar{I};\ \bar{\lambda}_i = 0 \},$$

$$I^* := J \cup I^+.$$

It may be convenient to define the so-called "critical cone" (or cone of critical directions):

$$C = \{ y \in \mathbb{R}^n;\ g'(\bar{x}) y \overset{\bar{I}}{\leqslant} 0;\ g'_{I^+}(\bar{x}) y = 0 \}.$$

Proposition 5.1. *Semistability of* (5.2) *is equivalent to*

$$\begin{cases} (y, \mu) = 0 \quad \textit{is the unique solution of} \\ H(\bar{x}, \bar{\lambda}) y + g'(\bar{x})^* \mu = 0, & (5.6\text{i}) \\ y \in C, \quad \mu_{I^0} \geq 0; \qquad \mu_i = 0 \quad \textit{if} \quad g_i(\bar{x}) < 0, \quad \forall i \in I; \qquad \mu_i g_i'(\bar{x}) y = 0, \quad \forall i \in I^0. & (5.6\text{ii}) \end{cases}$$

Proof. We have to prove the equivalence of (5.5) and (5.6). The set of solutions of (5.6i–ii) is a cone. Hence it is equivalent to state that $(y, \mu) = 0$ is the unique solution of (5.6i–ii) or to state that $(y, \mu) = 0$ is an isolated solution of (5.6i–ii)). Now it is sufficient to prove the equivalence of (5.5ii) and (5.6ii) when (y, μ) is small enough. If μ is sufficiently close to zero and $i \in I^+$, then $\bar{\lambda}_i + \mu_i > 0$, hence, by (5.5ii), $g_i'(\bar{x}) y = 0$. On the other hand if (5.5ii) holds, μ_{I^0} must be nonnegative and $\mu_i > 0$ for some $i \in I_0$ implies $g_i'(\bar{x}) y = 0$. Also if $g_i(\bar{x}) < 0$, then $g_i(\bar{x}) +$

$g_i'(\bar{x})y < 0$ for y sufficiently close to 0. For that reason (5.5ii) is equivalent (when (y, μ) is small enough) to

$$\begin{cases} g_i'(\bar{x})y = 0, & \forall i \in I^+, \\ g_i'(\bar{x})y \leq 0, \quad \mu_i \geq 0, \quad \mu_i g_i'(\bar{x})y = 0, & \forall i \in I^0, \\ \mu_i = 0 \quad \text{if} \quad g_i(\bar{x}) < 0, \end{cases}$$

and this is easily shown to be equivalent to (5.6ii). □

Let us now consider Newton's method applied to (5.2). The subproblem to be solved at step k is, denoting by d^k the increment in x, i.e., $d^k = x^{k+1} - x^k$,

$$\begin{cases} F(x^k) + H(x^k, \lambda^k)d^k + g'(x^k)^*\lambda^{k+1} = 0, \\ g(x^k) + g'(x^k)d^k \in N_1(\lambda^{k+1}). \end{cases}$$

As the evaluation of $g'(x^k)$ is already necessary in order to evaluate $\varphi(x^k, \lambda^k)$ the only part of the Jacobian that perhaps needs to be approximated is $H(x^k, \lambda^k)$. We then obtain the Newton-type algorithm:

Algorithm 3

0. Choose $(x^0, \lambda^0) \in \mathbb{R}^n \times \mathbb{R}^p$; $k \leftarrow 0$.
1. While (x^k, λ^k) is not a solution of (5.2): choose M^k, an $n \times n$ matrix, compute the (d^k, λ^{k+1}) solution of

$$\begin{cases} F(x^k) + M^k d^k + g'(x^k)^*\lambda^{k+1} = 0, \\ g(x^k) + g'(x^k)d^k \in N_1(\lambda^{k+1}), \end{cases}$$

and put $x^{k+1} \leftarrow x^k + d^k$.

When $M^k = H(x^k, \lambda^k)$, by applying Corollary 2.1 and Proposition 5.1, we easily obtain

Theorem 5.1 (Convergence of Newton's Method). *Let $\{x^k, \lambda^k\}$ be computed by Algorithm* 3 *with $M^k = H(x^k, \lambda^k)$ converging toward $(\bar{x}, \bar{\lambda})$ satisfying* (5.2) *and* (5.6). *If $x \to (F(x), g'(x))$ is C^1 (resp. C^1 with a locally Lipschitz derivative), then $(x^k, \lambda^k) \to (\bar{x}, \bar{\lambda})$ superlinearly (resp. at a quadratic rate).*

We now consider conditions related to the superlinear convergence of $\{x^k\}$ alone. We are looking for necessary and/or sufficient conditions of the following type: at each interation k we define

E^k is a closed convex subset of $\mathbb{R}^n$,
P^k is an orthogonal projection onto E^k,
$h^k := P^k[(H(\bar{x}, \bar{\lambda}) - M^k)d^k]$.

The condition will be

$$h^k = o(d^k). \tag{5.7}$$

As a particular case of our results we recover the characterization of Boggs *et al.* [3] concerning equality-constrained nonlinear programming problems, and we are able to extend the characterization to variational inequalities satisfying the assumption

$$d^t H(\bar{x}, \bar{\lambda}) d > 0 \qquad \text{for all } d \text{ in } C - \{0\}. \tag{5.8}$$

All our results, however, need the following qualification hypothesis (linear independence of the gradients of active constraints):

$$\{\nabla g_i(\bar{x})\}_{i \in \bar{I} \cup J} \text{ surjective.} \tag{5.9}$$

On the other hand we do not need any strict complementary hypothesis.

Theorem 5.2. *We assume $x \to (F(x), g'(x))$ to be C^1. Let $\{(x^k, \lambda^k)\}$ be computed by Algorithm* 3 *converging toward $(\bar{x}, \bar{\lambda})$, the semistable solution of* (5.2) *satisfying* (5.9). *Then*:

(i) *Condition* (5.7) *is sufficient for superlinear convergence when E^k is defined as*

$$E_1^k := \{d \in \ker g'_{I^*}(x^k);\ g'_i(x^k) d \geq 0,\ \forall i \in I^0 \text{ such that } g_i(x^k) + g'_i(x^k) d^k = 0\}.$$

(ii) *Condition* (5.7) *is necessary, and also sufficient, for superlinear convergence if, in addition,* (5.8) *holds, when E^k is defined as*

$$E_2^k := \{d \in \ker g'_{I^*}(x^k);\ g'_{I^0}(x^k) d \leq 0\}.$$

Remark 5.1. If the strict complementarity hypothesis holds, i.e., $I^0 = \emptyset$, then $E_1^k = E_2^k = \ker g'_{I^*}(x^k)$ and with this choice of E^k, condition (5.7) is necessary and sufficient for superlinear convergence of $\{x^k\}$, under the hypotheses of the semistability of $(\bar{x}, \bar{\lambda})$.

Proof of Theorem 5.2. (a) Preliminaries. Writing the Kuhn–Tucker conditions for the projection problem defining h^k we get the existence of $\eta^k \in \mathbb{R}^p$ satisfying

$$h^k - [H(\bar{x}, \bar{\lambda}) - M^k] d^k + g'(x^k)^* \eta^k = 0 \tag{5.10}$$

and

$$\eta_i^k = 0 \qquad \text{if} \quad i \in I - \bar{I},$$

$$\eta_{I^0}^k \leq 0, \quad \eta_i^k = 0 \qquad \text{for all } i \text{ in } I^0 \text{ such that} \quad g_i(x^k) + g'_i(x^k) d^k < 0, \quad \text{if} \quad E^k = E_1^k,$$

$$\eta_{I^0}^k \geq 0 \qquad \text{if} \quad E^k = E_2^k.$$

Subtracting the first relation defining the Newton-type step from (5.10) we get

$$h^k - F(x^k) - H(\bar{x}, \bar{\lambda}) d^k + g'(x^k)^* (\eta^k - \lambda^{k+1}) = 0. \tag{5.11}$$

Expanding $F(x^k)$ up to first order and taking (5.2) in account we have

$$\begin{aligned}-F(x^k) &= -F(\bar{x}) - F'(\bar{x})(x^k - \bar{x}) + o(x^k - \bar{x})\\ &= g'(\bar{x})^*\bar{\lambda} - F'(\bar{x})(x^k - \bar{x}) + o(x^k - \bar{x})\\ &= g'(x^k)^*\bar{\lambda} - H(\bar{x}, \bar{\lambda})(x^k - \bar{x}) + o(x^k - \bar{x}),\end{aligned}$$

hence, with (5.11),

$$h^k - H(\bar{x}, \bar{\lambda})(x^k + d^k - \bar{x}) + g'(x^k)^*(\bar{\lambda} + \eta^k - \lambda^{k+1}) = o(x^k - \bar{x}). \tag{5.12}$$

Let us define

$$\delta^k := \|h^k\| + \|x^k - \bar{x}\| + \|d^k\|.$$

Then $(\delta^k)^{-1}(h^k, x^k + d^k - \bar{x}, \lambda^{k+1} - \eta^k - \bar{\lambda})$ is bounded, the boundedness of the third term being a consequence of (5.9) and (5.12). Let (h, z, ζ) be a limit-point of this sequence, i.e., a limit for a subsequence $k \in S \subset \mathbb{N}$. Then $\zeta_i = 0$ if $i \in I - \bar{I}$ and from (5.12) we deduce that

$$h - H(\bar{x}, \bar{\lambda})z - g'(\bar{x})^*\zeta = 0. \tag{5.13}$$

Also, expanding g as follows,

$$g(x^k) + g'(x^k)d^k = g(\bar{x}) + g'(\bar{x})(x^k + d^k - \bar{x}) + o(x^k - \bar{x}),$$

we deduce from the fact that d^k is a Newton-type step associated with a multiplier λ^{k+1} that

$$\begin{cases} g'(\bar{x})z \overset{\bar{I}}{\leqslant} 0, & \\ g_i'(\bar{x})z = 0 & \text{if} \quad i \in I \quad \text{is such that} \quad \lambda_i^{k+1} > 0 \quad \text{for all} \quad k \in S \\ & \text{(which is the case if } \bar{\lambda}_i > 0, \text{ i.e., } i \in I^+). \end{cases} \tag{5.14}$$

The above relations imply

$$z \in C. \tag{5.15}$$

We also have, from the definition of h^k,

$$g'_{I^*}(\bar{x})h = 0, \tag{5.16}$$

$$h \in C \qquad \text{when} \quad E^k = E_2^k. \tag{5.17}$$

(b) Proof of case (i). If $h^k = o(d^k)$, then *a fortiori* $h = 0$. With (5.13) we deduce that

$$H(\bar{x}, \bar{\lambda})z + g'(\bar{x})^*\zeta = 0. \tag{5.18}$$

If $i \in I^0$ is such that $g_i'(\bar{x})z < 0$, then $g_i(x^k) + g_i'(x^k)d^k < 0$, $\eta_i^k = 0$, and $\lambda_i^{k+1} = 0$ for k in S large enough, hence $\zeta_i = 0$. This implies

$$\zeta_i g_i'(\bar{x})z = 0 \qquad \text{for all } i \text{ in } I^0.$$

However, this with the fact that $\zeta_{I^0} \geq 0$ (due to $\eta^k_{I^0} \leq 0$), (5.15), (5.18), and (5.6) imply $z = 0$, i.e.,

$$x^{k+1} - \bar{x} = o(\|h^k\| + \|x^k - \bar{x}\| + \|d^k\|).$$

This and (5.7) imply

$$x^{k+1} - \bar{x} = o(\|x^k - \bar{x}\| + \|x^{k+1} - x^k\|) = o(\|x^k - \bar{x}\| + \|x^{k+1} - \bar{x}\|),$$

which implies $x^{k+1} - \bar{x} = o(x^k - \bar{x})$, i.e., x^k converges superlinearly.

(c) Proof of case (ii). If x^k converges superlinearly, then $z = 0$. Computing the scalar product of (5.13) by h we get

$$\|h\|^2 = \langle \zeta, g'(\bar{x})h \rangle.$$

Using (5.17), the nonnegativity of $\lambda^{k+1}_{I^0}$, and the complementarity condition $\eta^k_i g'_i(x^k)h^k = 0$, for i in I^0, we deduce that the right-hand side of the above relation is nonpositive; hence $h = 0$, i.e.,

$$h^k = o(\|h^k\| + \|x^k - \bar{x}\| + \|d^k\|),$$

which implies $h^k = o(\|x^k - \bar{x}\| + \|d^k\|)$. However, the superlinear convergence of $\{x^k\}$ implies $\|d^k\|/\|x^k - \bar{x}\| \to 1$, hence (5.7) holds.

We now prove that if (5.7) and (5.8) hold, $\{x^k\}$ converges superlinearly. As (5.7) implies $h = 0$, computing the scalar product of (5.13) with z we get

$$\langle z, H(\bar{x}, \bar{\lambda})z \rangle + \langle \zeta, g'(\bar{x})z \rangle = 0. \tag{5.19}$$

As $z \in C$, $g'_i(\bar{x})z = 0$ if $\lambda^{k+1}_i \neq 0$, and $\eta^k_{I^0} \geq 0$, the second term of (5.19) is nonnegative. This and (5.8) imply that $z = 0$, i.e.,

$$x^{k+1} - \bar{x} = o(\|h^k\| + \|x^k - \bar{x}\| + \|d^k\|).$$

Using (5.7) and the relation $\|d^k\| \leq \|x^{k+1} - \bar{x}\| + \|x^k - \bar{x}\|$ we deduce that

$$x^{k+1} - \bar{x} = o(\|x^k - \bar{x}\| + \|x^{k+1} - \bar{x}\|),$$

which implies the superlinear convergence of $\{x^k\}$. $\square$

With the help of Theorem 5.2 we may obtain the superlinear convergence of $\{x^k\}$ when M^k is updated using ideas of quasi-Newton algorithms. We define the quasi-Newton equation (for M^{k+1}) as follows:

$$M(x^{k+1} - x^k) = F(x^{k+1}) - F(x^k) + [g'(x^{k+1}) - g'(x^k)]^* \lambda^{k+1}. \tag{5.20}$$

We assume that a closed convex subset $\mathscr{K}$ of the space of $n \times n$ matrices exists such that

$$H(x, \lambda) \in \mathscr{K}, \qquad \forall (x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^p, \tag{5.21}$$

and we choose the M^{k+1} solution of

$$\min \|M - M^k\|_\#; \qquad M \in \mathscr{K}; \qquad M \text{ satisfies (5.20)}, \tag{5.22}$$

where as before $\|\cdot\|_\#$ is a norm associated to a sclar product.

Lemma 5.1. *Under the hypotheses of Theorem 5.2, if M^{k+1} satisfies (5.20) and*

$$(M^{k+1} - M^k)(x^{k+1} - x^k) = o(x^{k+1} - x^k), \tag{5.23}$$

then $\{x^k\}$ converges superlinearly.

Proof. As M^{k+1} satisfies (5.20) we have

$$\begin{aligned} M^{k+1}(x^{k+1} - x^k) &= F(x^{k+1}) - F(x^k) + (g'(x^{k+1}) - g'(x^k))^*\bar{\lambda} \\ &\quad + (g'(x^{k+1}) - g'(x^k))^*(\lambda^{k+1} - \bar{\lambda}) \\ &= H(\bar{x}, \bar{\lambda})(x^{k+1} - x^k) + o(x^{k+1} - x^k). \end{aligned}$$

Hence if M^{k+1} satisfies (5.20), and (5.23) holds, then

$$(H(\bar{x}, \bar{\lambda}) - M^k)(x^{k+1} - x^k) = o(x^{k+1} - x^k).$$

Now let h^k be the projection of $(H(\bar{x}, \bar{\lambda}) - M^k)(x^{k+1} - x^k)$ onto E_1^k. As E_1^k is a cone, and the projector operator is nonexpansive we obtain

$$\|h^k\| \le \|(H(\bar{x}, \bar{\lambda}) - M^k)(x^{k+1} - x^k)\| = o(x^{k+1} - x^k).$$

This and Theorem 5.2 (case (i)) imply the conclusion. □

Theorem 5.3. *Let $(\bar{x}, \bar{\lambda})$ be a semistable and hemistable solution of (5.1). Then $\varepsilon > 0$ exists such that if $\|x^0 - \bar{x}\| + \|M^0 - H(\bar{x}, \bar{\lambda})\|_\# < \varepsilon$, then at each step k, a (x^{k+1}, λ^{k+1}) solution of the Newton-type step satisfying $\|x^{k+1} - x^k\| < 2\varepsilon$ exists. The sequence $\{x^k\}$ defined in this way converges superlinearly toward $\bar{x}$.*

Proof. Define

$$S^k := \{M \in \mathscr{K}; M \text{ satisfies } (5.20)\},$$

$$A^k := \int_0^1 H(x^k + \sigma(x^{k+1} - x^k), \lambda^{k+1})\, d\sigma.$$

Then A^k is an element of S^k and, for some $c_1 > 0$,

$$\|A^k - H(\bar{x}, \bar{\lambda})\|_\# \le c_1 v^k, \tag{5.24}$$

with

$$v^k := \|x^{k+1} - \bar{x}\| + \|x^k - \bar{x}\| + \|\lambda^{k+1} - \bar{\lambda}\|.$$

As M^{k+1} is the projection of M^k onto S^k, we have, from [12],

$$\|M^{k+1} - M^k\|_\#^2 + \|M^{k+1} - A^k\|_\#^2 \le \|M^k - A^k\|_\#^2, \tag{5.25}$$

hence with (5.24)

$$\|M^{k+1} - H(\bar{x}, \bar{\lambda})\|_\# \le \|M^k - H(\bar{x}, \bar{\lambda})\|_\# + 2c_1 v^k. \tag{5.26}$$

As a consequence the approximation of the Jacobian at step k is

$$\hat{M}^k := \begin{pmatrix} M^k & g'(x^k)^* \\ g'(x^k) & 0 \end{pmatrix}$$

and approximates

$$\bar{M} := \begin{pmatrix} H(\bar{x}, \bar{\lambda}) & g'(\bar{x})^* \\ g'(\bar{x}) & 0 \end{pmatrix}.$$

We define a new norm as follows. To

$$\hat{M} := \begin{pmatrix} M_{11} & M_{12} \\ M_{12}^* & 0 \end{pmatrix}$$

is associated

$$\|\hat{M}\|_\$:= \|M_{11}\|_\# + \|M_{12}\|$$

with $\|\cdot\|$ an arbitrary norm. From (5.26) we deduce that, for some $c_2 > 0$,

$$\|\hat{M}^{k+1} - \bar{M}\|_\$ \leq \|\hat{M}^k - \bar{M}\|_\$ + c_2 \nu^k.$$

Applying Lemma 2.2 (for which we may assume that $\lambda^0 = \bar{\lambda}$) we deduce that if (x^0, M^0) is close enough to $(\bar{x}, H(\bar{x}, \bar{\lambda}))$, then (x^k, λ^k) is well defined and converges linearly to $(\bar{x}, \bar{\lambda})$ and that $\|\hat{M}^k - \bar{M}\|_\$$ converges. This implies that $\|M^k - H(\bar{x}, \bar{\lambda})\|_\#$ converges. As $A^k \to H(\bar{x}, \bar{\lambda})$, $\|M^{k+1} - A^k\|_\#$ and $\|M^k - A^k\|_\#$ converge to the same limit, and with (5.25) this implies $\|M^{k+1} - M^k\| \to 0$. The conclusion is then a consequence of Lemma 5.1. □

6. Application to Nonlinear Programming

In this section we particularize some of our results to nonlinear programming problems, and we see that it allows us to get some improvements with respect to known results. By a nonlinear programming problem we mean

$$\min f(x); \qquad g(x) \preccurlyeq 0, \tag{6.1}$$

where f is a smooth mapping $\mathbb{R}^n \to \mathbb{R}$, and g as well as the relation "$\preccurlyeq$" are as in Section 5. Let us recall some well-known facts of optimization theory (see, e.g., [10]). To problem (6.1) is associated the first-order optimality system

$$\begin{cases} \nabla f(x) + g'(x)^* \lambda = 0, \\ g(x) \preccurlyeq 0, \qquad \lambda_I \geq 0, \qquad \lambda^t g(x) = 0, \end{cases} \tag{6.2}$$

which is formally equivalent to (5.1) if we define $F(x) := \nabla f(x)$. In this case the mapping $H(x, \lambda)$ can be interpreted as the Hessian with respect to x of the Lagrangean $L(x, \lambda) := f(x) + \lambda^t g(x)$. We say that λ is a Lagrange multiplier associated with x if (x, λ) satisfies (6.2). We recall the results involving second-order conditions with a unique multiplier.

Proposition 6.1 *(see, e.g.,* [1]*).*

(i) *(Second-order necessary condition.) Let* $\bar{x}$ *be a local solution of* (6.1) *to which a unique multiplier* $\bar{\lambda}$ *is associated. Then* $d^t H(\bar{x}, \bar{\lambda})d \geq 0$ *for all critical directions* d.

(ii) *(Second-order sufficiency condition.) Let* $(\bar{x}, \bar{\lambda})$ *satisfying* (6.2) *be such that* $d^t H(\bar{x}, \bar{\lambda})d > 0$ *for all nonzero critical directions* d. *Then* $\bar{x}$ *is a local solution of* (6.1).

We now make the link between semistability and the second-order sufficiency condition.

Proposition 6.2. *Let* $(\bar{x}, \bar{\lambda})$ *be an isolated solution of* (6.2) *such that* $\bar{x}$ *is a local solution of* (6.1). *Then* $(\bar{x}, \bar{\lambda})$ *is semistable iff it satisfies the second-order sufficiency condition.*

Proof. Characterization (3.2) of semistability applied to the variational inequality in form (5.1), and using (5.4), gives

$$\langle d, H(\bar{x}, \bar{\lambda})d\rangle > 0 \qquad \text{for all} \quad (d, \mu) \neq 0 \quad \text{the solution of}$$

$$\lambda_I + \mu_I \geq 0,$$

$$H(\bar{x}, \bar{\lambda})d + g'(\bar{x})^* \mu = 0,$$

$$g(\bar{x})^t(\mu - \lambda) = 0,$$

$$g(\bar{x}) + g'(\bar{x})d \in N_1(\lambda),$$

the last relation implying that d is critical. Hence the second-order sufficiency optimality condition implies semistability. Conversely, let us assume that the second-order sufficiency condition does not hold. By Proposition 6.1 a critical direction $d \neq 0$ exists with $\bar{d}^t H(\bar{x}, \bar{\lambda})\bar{d} = 0$, and $\bar{d}$ is a solution of the quadratic homogeneous problem

$$\min \tfrac{1}{2} d^t H(\bar{x}, \bar{\lambda})d; \qquad d \in C,$$

where the critical cone is

$$C := \{d;\ g'(\bar{x})d \overset{I}{\leqslant} 0;\ g_i'(\bar{x})d = 0 \text{ if } \lambda_i > 0,\ \forall i \in I\},$$

Writing the optimality system of this problem, we find that a multiplier μ is associated to $\bar{d}$ such that $(\bar{d}, \mu)$ satisfies (5.6i–ii). By Proposition 5.1 this contradicts semistability. □

Proposition 6.3. *Let* $(\bar{x}, \bar{\lambda})$ *be a semistable solution of* (6.2) *such that* $\bar{x}$ *is a local solution of* (6.1). *Then* $(\bar{x}, \bar{\lambda})$ *is hemistable.*

Proof. Semistability implies the uniqueness of the multiplier, hence also the hypothesis of Mangasarian and Fromovitz [17]. By Proposition 6.2 the second-order sufficiency condition also holds for problem (6.1) at $(\bar{x}, \bar{\lambda})$. Let us consider the following problem:

$$\min_d \nabla f(\bar{x})^t d + \tfrac{1}{2} d^t H(\bar{x}, \bar{\lambda}) d; \qquad g(\bar{x}) + g'(\bar{x}) d \leqslant 0. \tag{6.3}$$

Obviously $\bar{d} = 0$ satisfies the first-order optimality condition associated with the unique multiplier $\bar{\lambda}$. Also $\bar{d} = 0$ satisfies the second-order sufficiency condition (their formulation for (6.3) at $(\bar{d} = 0, \bar{\lambda})$ coincides with the one for (6.1) at $(\bar{x}, \bar{\lambda})$). Hence if we make small perturbations in the data of this problem, then a local solution exists whose distance to $\bar{d} = 0$ is of the order of the perturbation (see, e.g., Theorem 4.1 of [20]). Hence hemistability holds. □

From Theorem 2.2 and Proposition 6.3 we deduce

Theorem 6.1. *Assume that f and g are C^2 with Lipschitz second derivatives, $\bar{x}$ is a local solution of (6.1), $\bar{\lambda}$ is the unique Lagrange multiplier associated with $\bar{x}$, and the second-order sufficiency condition holds. Then $\varepsilon > 0$ exists such that if $\|x^0 - \bar{x}\| + \|\lambda^0 - \bar{\lambda}\| < \varepsilon$ and (x^{k+1}, λ^{k+1}) is chosen so that*

$$\|x^{k+1} - x^k\| + \|\lambda^{k+1} - \lambda^k\| < 2\varepsilon,$$

then Algorithm 3 with $M^k = H(x^k, \lambda^k)$, i.e., Newton's method, is well defined and converges at a quadratic rate to $(\bar{x}, \bar{\lambda})$.

Remark 6.1. That Newton's method converges at a quadratic rate when the starting point is close to a solution $(\bar{x}, \bar{\lambda})$ of (6.2), assuming x is a local solution of (6.1), the gradients of active constraints linearly independent, and strict complementarity, is well known. Recently the author [4] relaxed the strict complementarity hypothesis. Here we improve the result of [4] by assuming that the multiplier is unique instead of the linear independence of the gradients of active constraints.

We now apply the results of Section 5 on the superlinear convergence of $\{x^k\}$ only. From Theorem 5.2 and the fact that condition (5.8) coincides with the second-order sufficient condition, we get:

Theorem 6.2. *Let $\bar{x}$ be a local solution of (6.1) such that the gradients of active constraints are linearly independent, let $\bar{\lambda}$ be a multiplier associated with $\bar{x}$, and let the second-order sufficient condition hold. If (x^k, λ^k) computed by Algorithm 3 converges to $(\bar{x}, \bar{\lambda})$, then $\{x^k\}$ converges superlinearly iff*

$$P^k[(H(\bar{x}, \bar{\lambda}) - M^k)d^k] = o(d^k),$$

with the P^k orthogonal projection on the set E_2^k defined as in Theorem 5.2.

Remark 6.2. If no inequality constraint is present, Theorem 6.2 reduces to a theorem of Boggs *et al.* [3]. Some necessary or sufficient conditions (but not the characterization given here) for problems with equalities and inequalities, without strict complementarity have been given by the author in [4].

Acknowledgment

The author thanks an anonymous referee for his useful remarks.

References

1. Ben-Tal A (1980) Second-order and related extremality conditions in nonlinear programming. J Optim Theory Appl 21:143–165
2. Bertsekas DP (1982) Constrained Optimization and Lagrange Multipliers Methods. Academic Press, New York
3. Boggs PT, Tolle JW, Wang P (1982) On the local convergence of quasi-Newton methods for constrained optimization. SIAM J Control Optim 20:161–171
4. Bonnans JF (1989) Local study of Newton-type algorithms for constrained problems. Proc 5th Franco–German Conf on Optimization, Dolecki S, ed. Lectures Notes in Mathematics, vol 1405. Springer-Verlag, Berlin, pp 13–24
5. Bonnans JF, Launay G (1992) An implicit trust region algorithm for constrained optimization. INRIA Report.
6. Brézis H (1973) Opérateurs maximaux monotones. North-Holland, Amsterdam
7. Broyden CG, Dennis JE, Moré JJ (1973) On the local and superlinear convergence of quasi-Newton methods. J Inst Math Applic 12:223–245
8. Dennis JE, Moré JJ (1974) A characterization of superlinear convergence and its application to quasi-Newton methods. Math Comput 28:549–560
9. Dennis JE, Moré JJ (1977) Quasi-Newton methods, motivation and theory. SIAM Rev 19:46–89
10. Fletcher R (1987) Practical Methods of Optimization, 2nd edn. Wiley, Chichester
11. Gabay D (1982) Application de la méthode des multiplicateurs aux inéquations variationnelles. In: Méthodes de lagrangien augmenté, Fortin M, Glowinski R, eds. Dunod, Paris, pp 279–308
12. Grzegòrski SM (1985) Orthogonal projections on convex sets for Newton-like methods, SIAM J Numer Anal 22:1208–1219
13. Harker PT, Pang JS (1990) Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications. Math Programming 48:161–220
14. Josephy NH (1979) Newton's method for generalized equations. Technical Summary Report No 1965, Mathematics Research Center, University of Wisconsin-Madison
15. Josephy NH (1979) Quasi-Newton methods for generalized equations. Summary Report No 1966, Mathematics Research Center, University of Wisconsin-Madison
16. Lions JL, Stampacchia G (1967) Variational inequalities. Comm Pure Appl Math 20:493–519
17. Mangasarian OL, Fromovitz S (1967) The Fritz John necessary optimality conditions in the presence of equality and inequality constraints. J Math Anal Appl 17:37–47
18. Reinoza A (1985) The strong positivity condition. Math Oper Res 10:54–62
19. Robinson SM (1980) Strongly regular generalized equations. Math Oper Res 5:43–62
20. Robinson SM (1982) Generalized equations and their applications, part II: application to nonlinear programming. Math Programming Stud 19:200–221
21. Robinson SM (1983) Generalized equations. In: Mathematical Programming, the State of the Art, Bachem A, Grötschel M, Korte B, eds. Springer-Verlag, New York, pp 346–367
22. Robinson SM (1987) Local structure of feasible sets in nonlinear programming, part III: stability and sensitivity. Math Programming Stud 30:45–66
23. Robinson SM (to appear) Newton's method for a class of nonsmooth functions
24. Shapiro A (1990) On concepts of directional differentiability. J Optim Theory Appl 66:477–487

Accepted 26 January 1993